YUE XIN

📱 (+86) 19821286629   🎨 Machine-Phoenix   ✉ xinyuexiong@sjtu.edu.cn   🔗 yuexin.netlify.app

SUMMARY

Research Interests: I have a keen interest in machine learning, interpretable AI, and large (language) models. Now I am focusing on reasoning models.

Experience: First author at top conferences; Internship experience at renowned internet companies; Top 5% in Grades; National Scholarship; Mathematics and Physics competitions

Skills: Solid mathematical background; over 4 years of research experience; over 6 years of programming experience.

EDUCATION

Shanghai Jiao Tong University (SJTU)   **June 2023 − Present**

M.S. in Information and Communication Engineering   *Shanghai, China*

- Advisor: Prof. Hongkai Xiong and Prof. Wenrui Dai

Shanghai Jiao Tong University (SJTU)   **Sep 2019 − June 2023**

B.S. in Electronic Science and Technology (major) and Computer Science and Technology (minor)   *Shanghai, China*

- Upon graduation (all courses):   Grades: 89.66/100   GPA: 3.85/4.3   Ranking: 4/56
  When receiving the National Scholarship (all courses): Grades: 90.10/100   GPA: 3.89/4.3   Ranking: 2/64
- A(90+) for all math/physics related courses(Calculus, Physics, Probability and Statistics, etc) and most major related courses(Signals and Systems, Progamming, Digital Image Processing, etc). A+(95+) for most major related experimental courses(Engineering Problem Modeling and Simulation, Engineering Practice and Technological Innovation).

PUBLICATIONS

- **SalaMAnder: Shapley-based Mathematical Expression Attribution and Metric for Chain-of-Thought Reasoning**
  Yue Xin, Chen Shen, Shaotian Yan, Yaoming Wang, Xiaosong Yuan, Chenxi Huang, Jieping Ye.
  EMNLP 2025 accepted.

- **GLEAM: Global Share Local Transform MoE for Downstream Transferring With Enhanced Parameter Efficiency**
  Jiarui Zhang, **Yue Xin**, Yaoming Wang, Wenrui Dai, Ziyang Zheng, Chenglin Li, Junni Zou, Hongkai Xiong.
  ECAI 2025 accepted.

- **Clarifying the Behavior and the Difficulty of Adversarial Training**
  Xu Cheng*, Hao Zhang*, **Yue Xin**, Wen Shen, Jie Ren, Quanshi Zhang.
  AAAI 2024 accepted. [paper]

- **Towards the Dynamics of a DNN Learning Symbolic Interactions**
  Qihan Ren, Yang Xu, Junpeng Zhang, **Yue Xin**, Dongrui Liu, Quanshi Zhang.
  NeurIPS 2024 accepted. [paper]

- **Enhancing Chain-of-Thought Reasoning with Critical Representation Fine-tuning**
  Chenxi Huang, Liang Xie, Chen Shen, Shaotian Yan, Sinan Fan, **Yue Xin**, Binbin Lin, Deng Cai, Jieping Ye.
  ACL 2025 (main) accepted. [paper]

- **Parameter-Efficient Cross-Layer Feature Fusion via Chebyshev Polynomial Unit**
  Yue Xin, Jiarui Zhang, Ziyang Zheng, Yaoming Wang, Wenrui Dai, Chenglin Li, Junni Zou, Hongkai Xiong.
  VCIP 2025 on submission

- **D2-RST: Dual-Dimensional Residual Side Tuning for Mitigating Feature Forgetting in Parameter-Efficient Transfer Learning**
  Yue Xin, Yaoming Wang, Wenrui Dai, Jiarui Zhang, Ziyang Zheng, Chenglin Li, Junni Zou, Hongkai Xiong.
  BMVC 2025 on submission

- **Generalizable Geometric Image Caption Synthesis**
  Yue Xin, Wenyuan Wang, Rui Pan, Bingxu Meng, Renjie Pi, Tong Zhang.
  NeurIPS 2025 on submission. [blog] [dataset] [code]

- **Towards Noise-Robust Medical Segmentation via Chebyshev-Attention-Based Semi-UNet**
  Yue Xin, Ziyang Zheng, Wenrui Dai, Chenglin Li, Junni Zou, Hongkai Xiong.
  WACV 2026 on submission

- **A Technical Report on LLM Distillation**
  Yue Xin, Shaotian Yan, Kaiyuan Liu, Rui Miao, Bing Wang, Sinan Fan, Chen Shen, Jieping Ye.
  Arxiv under preparation

- **Perplexity-Aware Pruning for Efficient Chain-of-Thought Distillation**
  Yue Xin, Jiaxin Huang.
  ICLR 2026 on submission

- **Bootstrap Prompt Learning with Feature Adaptation for Vision-Language Efficient Tuning**
  Jiarui Zhang, Yaoming Wang, **Yue Xin**, Wenrui Dai, Ziyang Zheng, Chenglin Li, Junni Zou, Hongkai Xiong.
  ICLR 2026 on submission

ACADEMIC RESEARCH EXPERIENCE

Institute of Media, Information and Network(min), SJTU   **Nov 2022 – Present**
Machine Learning and Computer Vision Intern and Master's Student   *Advisor: Hongkai Xiong, Wenrui Dai*

- Proposed Chebyshev Fusion Unit (CFU), a lightweight feature fusion method to mitigate the challenge of balancing expressiveness and efficiency. Specifically, CFU computed high-order Chebyshev polynomial terms between residual and current-layer features to explicitly model complex cross-layer dependencies with minimal parameters. Comprehensive experiments verify its strong approximation and optimization capability. (VCIP 2025 on submission)
- Proposed Dual-Dimensional Residual Side Tuning (D2-RST) framework to mitigate feature forgetting and progressive spectral decay in deep layers by employing a dual-block side-tuning structure with low-rank linear mapping on aggregated features, and introducing an additional spatial-dimension pathway in parallel with the feature-dimension pathway. The properties and performance of RST are verified through mathematical proof and various experiments. (BMVC 2025 on submission)
- Proposed Chebyshev-Attention-Based Semi-Unet (CASUNet), a noise-resilient framework integrating a Semi-UNet backbone with a novel CPA (Chebyshev Polynomial Aggregation) module by first aggregating hierarchical features then expanding to orthogonal polynomial terms. Theoretical and experimental analysis verify its superior noise immunity and competitive performance. (WACV 2026 on Submission)
- Proposed GLEAM, an efficient fine-tuning method for large model parameters. This method leverages the high similarity of parameter matrices in LoRA to construct a low-rank decomposition, further reducing the number of parameters required for fine-tuning while enhancing performance. (ECAI 2025 accepted)
- Proposed PAT, a fault-tolerant multimodal classifier to solve the problem of conflict between prompter learning and adapter tuning. Specifically, it aligns model representations obtained from Soft Prompt and Adapter-based methods and incorporates contrastive learning loss to enhance model performance and generalization. (ICLR 2026 on submission)

Hint Lab, WashU   **Jun 2025 – Present**
LLM Research Intern   *Advisor: Jiaxin Huang*

- Proposed a perplexity-based method to analyze the long CoT pattern for data selection, after which removed the steps with low perplexity to construct a new dataset. Then conducted SFT and RL on the dataset to enhance the reasoning capacity of LLMs more efficiently without significant degradation of the output diversity. (ICLR 2026 on submission)

Tong's Lab, UIUC   **Mar 2025 – August 2025**
MLLM Research Intern   *Advisor: Tong Zhang*

- Proposed Geo-Image-Textualization, a reinforcement learning-based framework to generate high-quality and geometry-centered multimodal data by utilizing a rule-based data generation pipeline and adopting RAFT to further optimize the captions. Then constructed a dataset named GeoReasoning-10K to bridge the gap between visual and linguistic modalities in the geometry domain. Extensive experiments verify the superiority of the dataset for improving multimodal reasoning capacity. (NeurIPS 2025 on submission)

Feitian Lab, Alibaba Cloud   **Mar 2024 – Mar 2025**
Interpretable LLM Research Intern   *Advisor: Jieping Ye*

- Proposed SalaMAnder, a Shapley-value-based framework for quantifying component-level contributions in CoT reasoning. Specifically, we develop an efficient stratified sampling algorithm to compute Shapley value for mathematical expression attribution and CoSP (Cardinality of Shapley Positives) metric. Theoretical derivation and comprehensive validation across multiple models and benchmarks present a robust monotonic correlation between CoSP and model performance, providing theoretical explanations for the empirical success of CoT. (EMNLP 2025 accepted)
- Proposed CRFT, a novel method that identifies and optimizes critical representations that integrate significant information from preceding layers or regulate subsequent layer representation. CRFT effectively optimizes the representations in a low-rank linear subspace through information flow analysis. (ACL 2025 accepted)
- Explored the mechanism of LLM distillation and explained the pattern of long CoT, after which significantly improved the distilled model's reasoning capacity.

Interpretable ML lab, SJTU   **Feb 2022 – Nov 2022**
Interpretable Machine Learning Intern   *Advisor: Quanshi Zhang*

- Theoretically derived the analytical solution for multi-step adversarial attacks, which explains the reasons behind the optimization difficulties in adversarial training. This is validated through experiments. (AAAI 2024 accepted)
- Theoretically derived the two-stage dynamic interaction process of DNNs, proving that the network learning process gradually encodes interactions of varying complexity. This provides a theoretical foundation for understanding overfitting. (NeurIPS 2024 accepted)

ACADEMIC COMPETITION (Selected)

The 20th Chinese Graduate Mathematical Modeling Competition:	*Nation level, Second Prize*	**2023**
The Mathematical Contest in Modeling:	*World level, Meritorious Winner (First Prize)*	**2021**
The Huawei Cloud 'Cloud Pioneers' Few-Shot Detection Competition:	*Nation level, Third Place*	**2021**
The 12th National College Student Mathematical Competition:	*City level, First Prize*	**2020**
The 2nd National 'August 1st Cup' Online Mathematics Competition:	*Nation level, Tenth Place*	**2020**
Chinese Physics Olympiad:	*Province level, First Prize*	**2018**

TECHNICAL SKILLS

Programming Languages: Proficient in Python, C++, Matlab, LaTeX, etc.

Frameworks: Proficient in PyTorch, NumPy, VLLM, SgLang, LLaMA-Factory.

Mathematics: Proficient in calculus, linear algebra, probability statistics, etc.

Language: mandarin (native), English (fluent)

HONORS & AWARDS (Selected)

Outstanding Undergraduate Graduate of Shanghai Jiao Tong University	*University level, 5%*	**2023**
National Scholarship	*Nation level, 2%*	**2021**
Shanghai Jiao Tong University A-Class Excellent Scholarship for Undergraduate	*University level, 2%*	**2021**
Shenzhen Stock Exchange Scholarship	*University level, 2%*	**2020**
Shanghai Jiao Tong University B-Class Excellent Scholarship for Undergraduate	*University level, 5%*	**2020**